



# Space-time estimation of a particle system model

Xavier Guyon, Besnik Pumo

## ► To cite this version:

Xavier Guyon, Besnik Pumo. Space-time estimation of a particle system model. *Statistics A Journal of Theoretical and Applied Statistics*, 2007, 41 (5), pp.395-407. 10.1080/02331880701399074 . hal-00256589

**HAL Id: hal-00256589**

**<https://hal.science/hal-00256589>**

Submitted on 8 Apr 2008

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Space-time estimation of a particle system model

Xavier Guyon<sup>a</sup>, Besnik Pumo<sup>b, \*,†</sup>

<sup>a</sup> SAMOS, Université Paris 1, 90 rue de Tolbiac, 75634 Paris 13e, France

<sup>b</sup> SAGAH, INH, 2 rue Le Nôtre, 49000 Angers, France

February 12, 2007

---

\*Corresponding author. Tel.: +02-41-22-54-96; fax.: +02-41-22-54-13

†E-mail addresses: Xavier.Guyon@univ-paris1.fr; Besnik.Pumo@inh.fr

## Abstract

Let  $X$  be a discrete time contact process (CP) on  $\mathbb{Z}^2$  as defined by Durrett and Levin (1994). We study the estimation of the model based on space-time evolution of  $X$ , that is,  $T + 1$  successive observations of  $X$  on a finite subset  $S$  of sites. We consider the maximum marginal pseudo-likelihood (MPL) estimator and show that, when  $T \rightarrow \infty$ , this estimator is consistent and asymptotically normal for a non vanishing supercritical CP. Numerical studies confirm these theoretical.

*Key words:* contact process, parameter estimation, supercritical process, sub-ergodicity, marginal pseudo-likelihood, triangular martingale array.

*AMS Classification:* 62M30, 62M05, 62F12

## 1 Introduction and description of the model

Consider a simple model of spread of a single species population evolving in  $\mathbb{Z}^2$ . Depending on some biological parameters, the dynamics is determined by specifying, for each site  $s \in \mathbb{Z}^2$ , the conditional probability that site  $s$  will be in state  $X_{t+1}(s) = y \in \{0, 1\}$  at time  $t + 1$  given  $X_t$ , the configuration at time  $t$ . State 1 (respectively 0) means that there is a (respectively no) plant in  $s$ . In this paper we propose an estimator for the parameters of the model, based on observations of  $X$  at instants  $t = 0, \dots, T$  on a finite and fixed subset  $S$  of  $\mathbb{Z}^2$  and study the asymptotic properties of the estimator when the process is non vanishing on  $S$ . Fiocco (1997) and Fiocco & Zwet (2003) considered the estimation problem based on one observation at time  $t$ , when  $t$  is sufficiently large.

We consider the *discrete time version* of the *Contact Process* (CP) as defined by Durrett & Levin (1994). So, let us suppose that the transition probability at a site  $s$  and

at time  $t$  is stationary in space and time and depends locally on  $x_{t-1}(\mathcal{N}_1(s))$ , the first order neighborhood of the site  $s$  at time  $t - 1$ , where  $\mathcal{N}_d(s) = \{u \in \mathbb{Z}^2 : \|s - u\|_1 \leq d\}$ .

Denote  $\partial s = \mathcal{N}_1(s) \setminus \{s\}$ . The system evolves as follows:

- a. Each plant alive at time  $t$  dies with a probability  $\gamma$  at time  $t + 1$ ,
- b. If a plant in  $s$  at time  $t$  survives, then for each  $u \in \partial s$  it produces an offspring that is sent to  $u$  at time  $t + 1$  with probability  $\lambda$  (a surviving plant in  $s$  may produce 0 to 4 offsprings in  $\partial s$ ); the reproducing events for different  $s$  and  $u \in \partial s$  are independent,
- c. If there is one or more offsprings sent in  $u$  at time  $t + 1$ , or if a plant in  $u$  at time  $t$  survives, then  $X_{t+1}(u) = 1$ ; otherwise  $X_{t+1}(u) = 0$ .

Furthermore, events defined on (a) and (b) are independent in space and time.

This model depends on the parameter  $\theta = (\gamma, \lambda)$  and we suppose that  $\theta \in ]0, 1[^2$ . Other models are possible by defining different rules of evolution (cf. Durrett & Levin, 1994; Griffeath, 1981). The methods developed here can be generalized to processes with non invariant transition in space and/or in time.

Denote  $\tau = \inf\{t : X_t(s) = 0 \text{ for each } s \in \mathbb{Z}^2\}$  the extinction time of a CP starting from a single occupied site. A CP is *supercritical* (Durrett, 1995) if there is a positive probability for the species surviving forever, that is,  $P(\tau = +\infty) > 0$ . Since the “all 0 in  $\mathbb{Z}^2$ ” is an absorbing state, we need a condition **(I)** of non-extinction of  $X$  on the domain of observation  $S$  to make sense for the asymptotic study. This condition, presented in §3, is verified for a non vanishing supercritical CP.

The paper is organized as follows. In section 2 we define the *marginal pseudo-likelihood* (MPL) estimator of  $\theta$  and in section 3 we give a result which shows that the MPL allows

to identify the parameter. The consistency and asymptotic normality of estimator are formulated in section 4. Proofs are postponed to section 6. Some numerical studies are presented in section 5.

## 2 Marginal pseudo-likelihood (MPL)

Let  $\mathbf{x}(T) = (x_0, x_1, \dots, x_T)$  be  $T + 1$  successive configurations of  $X$ ,  $S$  a finite subset of  $\mathbb{Z}^2$  and  $S_1 = \{u \in \mathbb{Z}^2 : \exists v \in S \text{ such that } \|u - v\|_1 \leq 1\}$ . The estimator of  $\theta$  we choose is a value which maximizes a Marginal Pseudo-Likelihood (MPL) of  $\mathbf{x}(T)$  observed on  $S_1$ . The idea of pseudo-likelihood is classic in statistic: gaussian pseudo-likelihood for stationary field on  $\mathbb{Z}^d$  (Whittle, 1954), conditional pseudo-likelihood for a Markov field on a lattice (Besag, 1974). Besag proposed a similar method for the estimation of parameters of a simple statistical model for the spread of disease over a rectangular lattice of plants (see the discussion of Besag in Mollison, 1977).

For  $A \subset S$ , and if  $x_t(B)$  is the configuration of  $X$  at time  $t$  on  $B$ , denote  $P_A(x_t, x_{t+1}; \theta)$  the transition-probability

$$P_A(x_t, x_{t+1}; \theta) = P(X_{t+1}(A) = x_{t+1}(A) \mid X_t(S_1) = x_t(S_1)).$$

As the transition-probability for  $A = S$  is analytically intractable, as  $\#(S)$ , the number of sites of  $S$ , is large, we use a *marginal pseudo-transition probability*  $M_S(x_t, x_{t+1}; \theta)$  on  $S$ , in order to estimate  $\theta$ .  $M_S(x_t, x_{t+1}; \theta)$  is the product of  $P_{\{s\}}(x_t, x_{t+1}; \theta)$  for  $s \in I(x_t, S)$ , where:

$$I(x_t, S) = \{s \in S \text{ such that } x_t(\mathcal{N}_1(s)) \neq 0\}$$

Since  $0 < P_{\{s\}}(x_t, x_{t+1}; \theta) < 1$  for any  $x_{t+1}(s)$  when  $s \in I(x_t, S)$  and  $P_{\{s\}}(x_t, x_{t+1}; \theta) = 0$  or 1 when  $x_t(\mathcal{N}_1(s)) = 0$  only sites of  $I(x_t, S)$  are informative for  $\theta$ . The product of these marginal pseudo-transitions at consecutive instants defines the MPL. For  $s \in S$  and  $A$  a finite subset of  $\mathbb{Z}^2$ , denote  $m(x_t, A) = \sum_{s \in A} x_t(s)$ , the number of sites of  $A$  occupied by  $x_t$ . As the model is isotropic in space, the law of  $X_{t+1}(s)$  given  $x_t$  depends only on  $c(x_t, s)$ , a summary for  $x_t(\mathcal{N}_1(s))$ , where :

$$c(x_t, s) = (x_t(s), m(x_t, \partial s)) \in \mathcal{C}_1 = \{0, 1\} \times \{0, 1, 2, 3, 4\}. \quad (1)$$

More precisely,  $X_{t+1}(s)$  conditionally to  $x_t$  is a Bernoulli random variable:

$$P_{\{s\}}(x_t, x_{t+1}; \theta) = p(x_t, s; \theta)^{1-x_{t+1}(s)}(1 - p(x_t, s; \theta))^{x_{t+1}(s)},$$

with parameter  $1 - p(x_t, s; \theta)$  where:

$$p(x_t, s; \theta) = \gamma^{x_t(s)} \delta^{m(x_t, \partial s)}. \quad (2)$$

$\delta = \gamma + (1 - \gamma)(1 - \lambda) = 1 - \lambda + \gamma\lambda$  controls non-proliferation at time  $t + 1$  in  $s$  from a plant present in  $s' \in \partial s$  at time  $t$ . Define

$$M_S(x_t, x_{t+1}; \theta) = \prod_{s \in I(x_t, S)} p(x_t, s; \theta)^{1-x_{t+1}(s)}(1 - p(x_t, s; \theta))^{x_{t+1}(s)} \quad (3)$$

if  $I(x_t, S) \neq \emptyset$ , and  $M(0, 0; \theta) = 1$  otherwise.

The following property that states the dependency between the conditional variables on different sites, will be useful for the calculation of the variance of our estimator. Denote  $\eta = \gamma + (1 - \gamma)(1 - \lambda)^2$  the probability of non-proliferation at time  $(t + 1)$  in the set  $\{s, s'\}$  of a plant present in  $u \in \partial s \cap \partial s'$  at time  $t$ ,  $\partial s \nabla \partial s' = (\partial s \setminus \mathcal{N}_1(s')) \cup (\partial s' \setminus \mathcal{N}_1(s))$ ; then,

$$P(X_{t+1}(\{s, s'\}) = (0, 0) \mid x_t) = \gamma^{m(x_t, \{s, s'\})} \delta^{m(x_t, \partial s \nabla \partial s')} \eta^{m(x_t, \partial s \cap \partial s')}$$

It follows that for  $s \neq s'$ :

$$\text{Cov}(X_{t+1}(s), X_{t+1}(s') \mid x_t) = p(x_t, s; \theta) p(x_t, s'; \theta) [b(x_t, s, s'; \theta) - 1]$$

where:

$$b(x_t, s, s'; \theta) = \begin{cases} \delta^{-m(x_t, \{s, s'\})} & \text{if } s' \in \mathcal{N}_1(s) \setminus \{s\} \\ \delta^{-2m(x_t, \partial s \cap \partial s')} \eta^{m(x_t, \partial s \cap \partial s')} & \text{if } s' \in \mathcal{N}_2(s) \setminus \mathcal{N}_1(s) \\ 1 & \text{if } s' \notin \mathcal{N}_2(s). \end{cases} \quad (4)$$

In particular if  $s' \notin \mathcal{N}_2(s)$ ,  $(X_{t+1}(s) \mid x_t)$  and  $(X_{t+1}(s') \mid x_t)$  are independent.

Let us give the explicit expression of MPL based on observation of  $\mathbf{x}(T)$  on  $S_1$ . Denote  $n(x_t)$  the number of informative sites of the configuration  $x_t$  on  $S$ :

$$n(x_t) = \sharp(I(x_t, S)), \text{ and } n(x(T)) = \sum_{t=0}^{T-1} n(x_t).$$

For  $c \in \mathcal{C}_1$ , let  $n(x_t, c)$  be the number of sites  $s \in S$  with configuration  $c$ :

$$n(x_t, c) = \#\{s \in S : c(x_t, s) = c\}, \text{ and } n(x(T), c) = \sum_{t=0}^{T-1} n(x_t, c).$$

Then  $n(\mathbf{x}(T)) = \sum_{c \neq (0,0)} n(\mathbf{x}(T), c)$ . The normalized log-marginal pseudo-likelihood of  $\mathbf{x}(T)$  observed on  $S_1$  is:

$$l_{T,S}(\mathbf{x}(T); \theta) = \frac{1}{n(\mathbf{x}(T))} \sum_{t=0}^{T-1} \log M_S(x_t, x_{t+1}; \theta).$$

As  $S$  is fixed, we omit in most cases the subscript  $S$ ; thus  $I(x)$  or  $M(x_t, x_{t+1}; \theta)$  stand for  $I(x, S)$  or  $M_S(x_t, x_{t+1}; \theta)$ . In the same way, for a fixed  $\mathbf{x}(T)$ , we will write  $n(T)$  for  $n(\mathbf{x}(T))$ ,  $n(T, c)$  for  $n(\mathbf{x}(T), c)$  and  $l_T(\theta) \equiv l_{T,S}(\mathbf{x}(T); \theta)$ . Using (3) we have:

$$l_T(\theta) = \frac{1}{n(T)} \sum_{t=0}^{T-1} \sum_{s \in I(x_t)} \{[1 - x_{t+1}(s)] \log p(x_t, s; \theta) + x_{t+1}(s) \log[1 - p(x_t, s; \theta)]\}. \quad (5)$$

The maximum MPL estimator of  $\theta$  (or MPLE) is a value which maximizes the MPL,

$$\hat{\theta}_T = \arg \max_{\theta} l_T(\theta).$$

### 3 Marginal PL allows identification of $\theta$

In this section we show that Marginal PL allows identification of  $\theta$  under the condition

(I) of non-extinction of  $X$  on  $S$  :

$$(I) : I_{\infty} = \{\mathbf{x} = (x_t, t \geq 0) \text{ such that } n(\mathbf{x}(T)) \rightarrow \infty \text{ as } T \rightarrow \infty\}.$$



In order to obtain this result we need to show that  $\pi_c$  are strictly positive for two linearly independent configurations, where:

$$\pi_c = \varliminf_{T \rightarrow \infty} \frac{n(T, c)}{n(T)}.$$

Let  $\mathcal{C}_1^*$  be the set of configurations on  $\mathcal{N}_1(0)$  such that  $x(0) = 1$ . The following Lemma states the positivity  $\pi_c > 0$  for  $c \in \mathcal{C}_1^*$  under the condition **(I)** and is a particular case of a *the subergodicity result* (see Lemma 4 in section 6) useful to prove positivity of some pseudo-information matrix.

**Lemma 1** *There exists  $\alpha > 0$  such that,  $\forall c \in \mathcal{C}_1^*$ , and  $\forall x \in I_\infty$ , we have  $\pi_c \geq \alpha$ .*

Under **(I)** and for large  $T$ ,  $\theta \rightarrow l_T(\theta)$  allows identification of  $\theta$ . Indeed:

- if  $x(T)$  realizes two linearly independent configurations  $c_a = (u_a, v_a)$  and  $c_b = (u_b, v_b)$ , then  $\theta \mapsto l_T(\theta)$  is an injective function;
- under **(I)**, the probability that each configuration  $c$  of  $\mathcal{C}_1^*$  appears on  $S$  converges to 1 when  $T \rightarrow \infty$ .

Remarks :

1. As a non-vanishing supercritical process is ergodic, with a spatially translation-invariant limite law  $\mu$  (Durrett, 1995) satisfying  $\mu_{\{0\}}(1) > 0$ , it follows that  $\pi_c = \lim_{T \rightarrow \infty} \frac{n(T, c)}{n(T)}$  exists and is strictly positive.
2. Space and/or time invariance of the transitions model is not crucial in the proof of the subergodicity result: a similar result can be proved for non translation invariant

models under some convenient condition that transition probabilities are uniformly positive.

## 4 Consistency and normality of the MPL estimator

### 4.1 The results

Let  $f : U \rightarrow \mathbb{R}$  be a real function twice continuously differentiable on an open subset  $U$  of  $\mathbb{R}^d$  and  $f^{(1)}(\theta)$  the vector of first derivatives. The following result sets up the consistency and asymptotic normality of the maximum MPLE  $\hat{\theta}_T$  associated with (5). The proof is postponed to section 6.

Let  $q(c; \theta) = P(X_{t+1}(s) = 0 \mid c(x_t, s) = c)$ , and  $A_T(\theta_o)$ ,  $B_T(\theta_o)$  the  $2 \times 2$  matrices:

$$A_T(\theta_o) = \frac{1}{n(T)} \sum_{t=0}^{T-1} \sum_{s \in I(x_t)} \frac{p^{(1)t}[p^{(1)}]}{p(1-p)}(x_t, s; \theta_o) \quad (6)$$

$$= \sum_{c \neq (0,0)} \frac{n(T, c)}{n(T)} \times \frac{q^{(1)t}[q^{(1)}]}{q(1-q)}(c; \theta_o)$$

$$B_T(\theta_o) = \frac{1}{n(T)} \sum_{t=0}^{T-1} \sum_{s \neq s' \in I(x_t)} [b(x_t, s, s'; \theta_o) - 1] \frac{p^{(1)}(x_t, s; \theta_o)^t [p^{(1)}(x_t, s'; \theta_o)]}{[1-p(x_t, s; \theta_o)] [1-p(x_t, s'; \theta_o)]} \quad (7)$$

with  $p(x_t, s; \theta)$  given by (2) and  $b(x_t, s, s'; \theta_o)$  by (4).

**Theorem 1** *Let  $\theta_o = (\gamma_o, \lambda_o)$  be an interior point of a compact  $\Theta \subset ]0, 1]^2$  and the true unknown value of the parameter  $\theta$ . Under condition (I):*

(1) *The maximum MPL estimator is consistent:*

$$\lim_{T \rightarrow \infty} \hat{\theta}_T \stackrel{a.s.}{=} \theta_o.$$

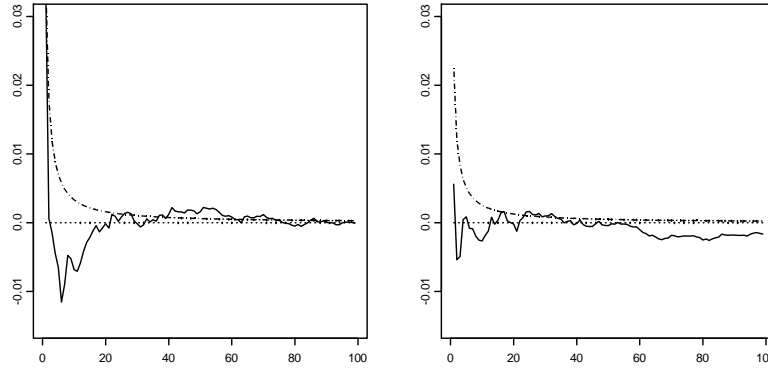


Figure 1: Evolution of the bias (solid lines) and standard deviation (multiplied by 100, dotted lines) for the estimators of  $\gamma_o$  (left) and  $\lambda_o$  (right) for the supercritical CP with parameters  $\gamma_o = 0.35, \lambda_o = 0.25$ .

(2) *This estimator is asymptotically normal,*

$$\sqrt{n(T)} [A_T(\theta_o) + B_T(\theta_o)]^{-1/2} A_T(\theta_o)(\hat{\theta}_T - \theta_o) \xrightarrow{d} \mathbf{G}_2(0, I_2)$$

where  $I_2$  is the  $2 \times 2$  identity matrix.

## 5 Numerical studies

In this section we give some empirical results with  $S$  the  $64 \times 64$  square lattice and initial configuration ‘all sites occupied’. To avoid boundary effects we have used periodic boundary conditions.

In Fig. 1 we present the evolution of the bias and the standard deviation of  $\hat{\gamma}_T$  and

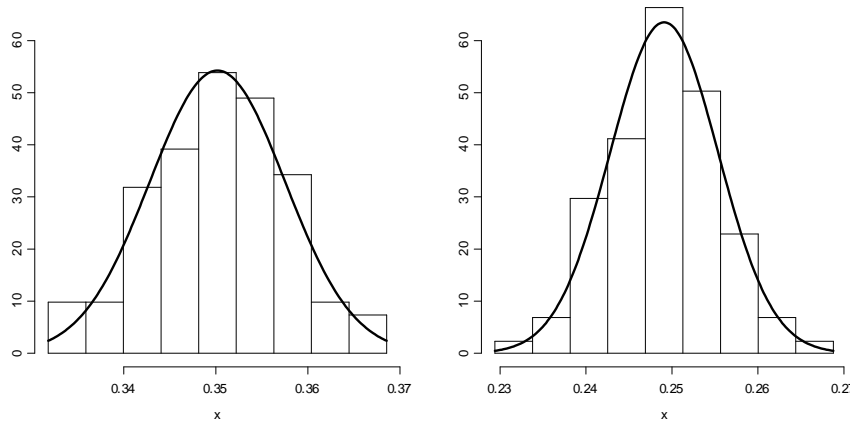


Figure 2: Histograms of 100 estimations of  $\gamma_o$  (left) and  $\lambda_o$  (right) for the supercritical CP with parameters  $\gamma_o = 0.35, \lambda_o = 0.25$ .

$\hat{\lambda}_T$  for  $T = 1, \dots, 99$  for the supercritical CP with parameters  $\gamma_o = 0.35, \lambda_o = 0.25$ .

Empirical study of asymptotic normality of estimators for a supercritical CP is based on 100 simulations with  $T = 99$ . Histograms are presented in Fig. 2. Asymptotic normality is checked by using a chi-squared test at level 5% and defining 9 equiprobable classes. Normality is accepted for  $\hat{\gamma}$  (resp.  $\hat{\lambda}$ ) since  $\chi^2 = 1.7$  (resp.  $\chi^2 = 4.4$ ) and  $\chi_{0.95}^2(6) = 12.59$ .

We also compared the estimated theoretical standard errors  $\hat{\sigma}_{\hat{\gamma}}, \hat{\sigma}_{\hat{\lambda}}$  and empirical standard errors  $s_{\hat{\gamma}}, s_{\hat{\lambda}}$  for the supercritical CP with parameter  $\gamma_o = 0.35, \lambda_o = 0.25$ . The values  $\hat{\sigma}_{\hat{\gamma}_4}, \hat{\sigma}_{\hat{\lambda}_4}$  are obtained from a single simulation with  $T = 4$  by applying Theorem 1 where  $A_4(\theta_o)$  (resp.  $B_4(\theta_o)$ ) are approximated by  $A_4(\hat{\theta}_4)$  (resp.  $B_4(\hat{\theta}_4)$ ). The empirical standard errors  $s_{\hat{\gamma}_4}, s_{\hat{\lambda}_4}$  are empirical estimations obtained from the 100 estimations associated to

the 100 simulations. The results are presented in table 1. We note that, as expected, there are few differences between estimated theoretical standard errors and empirical standard errors.

	$\hat{\sigma}_{\hat{\gamma}_4}$	$s_{\hat{\gamma}_4}$	$\hat{\sigma}_{\hat{\lambda}_4}$	$s_{\hat{\lambda}_4}$
Estimations	0.0076	0.0079	0.0059	0.0064

Table 1: Theoretical and empirical standard errors of the estimators.

Finally, Table 2 gives the estimations of  $\gamma_o$  and  $\lambda_o$  for twelve CP with parameters  $(\gamma_o, \lambda_o) \in (0.2, 0.4, 0.6) \times (0.1, 0.2, 0.3, 0.4)$ . In these simulations,  $T = 4$  and 40% of randomly chosen sites were occupied at time  $t = 0$ .

$\gamma$	$\lambda = 0.1$					$\lambda = 0.2$				
	$\hat{\gamma}_4$	$\hat{\sigma}_{\hat{\gamma}_4}$	$\hat{\lambda}_4$	$\hat{\sigma}_{\hat{\lambda}_4}$	$n(4)$	$\hat{\gamma}_4$	$\hat{\sigma}_{\hat{\gamma}_4}$	$\hat{\lambda}_4$	$\hat{\sigma}_{\hat{\lambda}_4}$	$n(4)$
0.2	0.210	0.006	0.104	0.003	14598	0.189	0.006	0.193	0.004	15286
0.4	0.391	0.008	0.106	0.004	12402	0.399	0.008	0.188	0.005	13565
0.6	0.597	0.009	0.100	0.005	9224	0.607	0.009	0.206	0.008	10452
$\gamma$	$\lambda = 0.3$					$\lambda = 0.4$				
	$\hat{\gamma}_4$	$\hat{\sigma}_{\hat{\gamma}_4}$	$\hat{\lambda}_4$	$\hat{\sigma}_{\hat{\lambda}_4}$	$n(4)$	$\hat{\gamma}_4$	$\hat{\sigma}_{\hat{\gamma}_4}$	$\hat{\lambda}_4$	$\hat{\sigma}_{\hat{\lambda}_4}$	$n(4)$
0.2	0.202	0.006	0.294	0.006	15761	0.198	0.008	0.400	0.008	15843
0.4	0.394	0.008	0.305	0.007	14729	0.393	0.009	0.400	0.009	15106
0.6	0.596	0.009	0.292	0.010	11468	0.604	0.009	0.386	0.012	12402

Table 2: Estimation of the parameters and their standard deviation.

## 6 Proofs of Theorem 1

### 6.1 Proof of the consistency of $\hat{\theta}_T$

Denote  $n_1(x_t, c)$  the number of occupied sites  $s \in S$  at time  $t + 1$  with configuration  $c(x_t, s) = c$  at time  $t$ ,  $I^A(x_t) = I(x_t) \cap A$ ,  $n^A(x_t) = \sharp(I^A(x_t))$  and  $n_1(T, c) = \sum_{t=0}^{T-1} n_1(x_t, c)$ .

**Lemma 2** For each subset  $A \subset S$ :  $Var \left[ \sum_{s \in I^A(x_t)} X_{t+1}(s) \mid x_t \right] \leq (13/4) \cdot n^A(x_t)$ .

*Proof* : This follows as  $(X_{t+1}(s) \mid x_t)$ ,  $(X_{t+1}(s') \mid x_t)$  are independent Bernoulli variables when  $s' \notin \mathcal{N}_2(s)$ ,  $\sharp(\mathcal{N}_2(s)) = 13$  and  $Var[X_{t+1}(s) \mid x_t] \leq 1/4$ . ..... ■

**Lemma 3** Let  $c \in \mathcal{C}_1$ . If  $n(T, c) \rightarrow \infty$  when  $T \rightarrow \infty$ , then  $\frac{n_1(T, c)}{n(T, c)} \xrightarrow{a.s.} 1 - q(c; \theta_o)$ .

*Proof* : We apply the following result: if  $(a_n)$  is a sequence of real positive numbers and  $b_n = \sum_{k=1}^n a_k$  then  $\sum_{n=1}^{\infty} \frac{a_n}{b_n^2} < \infty$ . For  $I(x, c) = \{s : s \in I(x), c(x, s) = c\}$ , let:

$$\begin{aligned} \zeta_t(c; \theta_o) &= \sum_{s \in I(x_{t-1}, c)} \{X_t(s) - [1 - q(c; \theta_o)]\} \text{ for } 1 \leq t \leq T \\ S_T(c; \theta_o) &= \sum_{t=1, T} \zeta_t(c; \theta_o), \text{ and } \mathcal{F}_T = \sigma(X_t, t \leq T) \end{aligned}$$

$(S_T(c; \theta_o))$  is a  $(\mathcal{F}_T)$ -martingale. To prove the announced result, we apply Theorem 2.18 of Hall & Heyde (1980) with  $X_T = \zeta_T(c; \theta_o)$ ,  $U_T = n(T, c)$ . It is sufficient to show that:

$$\sum_{T=1}^{\infty} \frac{E_{\theta_o} ([\zeta_T(c; \theta_o)]^2 \mid x_{T-1})}{[n(T, c)]^2} < \infty.$$

This follows from lemma 2 since  $E_{\theta_o} ([\zeta_T(c; \theta_o)]^2 \mid x_{T-1}) \leq (13/4) \cdot n(x_{T-1}, c)$  ..... ■

To prove the consistency of  $\hat{\theta}_T$ , we apply Theorem 3.4.3 of Guyon (1995; see also Dacunha-Castelle D. and Duflo, M., 1986). Let:

$$\begin{aligned} k_c(\theta_o, \theta) &= [1 - q(c; \theta_o)] \left\{ \log \frac{1 - q(c; \theta_o)}{1 - q(c; \theta)} - \log \frac{q(c; \theta_o)}{q(c; \theta)} \right\} + \log \frac{q(c; \theta_o)}{q(c; \theta)}, \\ K(\theta_o, \theta) &= \sum_{c \neq (0,0)} \pi_c \times k_c(\theta_o, \theta). \end{aligned}$$

$K(\theta_o, \theta)$  have a single minimum at  $\theta_o$ : indeed,  $K(\theta_o, \theta_o) = 0$ , and  $K(\theta_o, \theta) > 0$  if  $\theta \neq \theta_o$  since  $k_c(\theta_o, \theta) > 0$  and from Theorem 1,  $\pi_c > 0$  for  $c \in \mathcal{C}_1^*$ .

Denote  $U_T(\theta) = -l_T(\theta)$ , the opposite of log-MPL. Then:

$$U_T(\theta) = - \sum_{c \neq (0,0)} \frac{n(T, c)}{n(T)} \left\{ \frac{n_1(T, c)}{n(T, c)} \cdot \log \frac{1 - q(c; \theta)}{q(c; \theta)} + \log[q(c; \theta)] \right\}$$

Since  $\frac{n_1(T, c)}{n(T, c)} \xrightarrow{as} 1 - q(c; \theta_o)$ , it follows that:

$$\underline{\lim}_{T \rightarrow \infty} [U_T(\theta) - U_T(\theta_o)] \geq K(\theta_o, \theta) \text{ a.s. on } I_\infty.$$

It remains to prove that there exists  $\epsilon_k \rightarrow 0$  such that  $P_o(\overline{\lim}_{T \rightarrow \infty} (W_T(1/k) \leq \epsilon_k)) = 1$  where  $W_T(\cdot)$  is the modulus of continuity of  $U_T(\theta)$ . This is a direct consequence of the compacity of  $\Theta \subseteq ]0, 1]^2$  and the fact that  $q$  is continuously differentiable. This completes the proof. .... ■

## 6.2 Proof of asymptotic normality of $\hat{\theta}_T$

We apply Theorem 3.4.5 of Guyon (1995) by verifying conditions (H1)-(H3) for the asymptotic normality of the minimum  $U_n$ -contrast consistent estimator :

(H1) There exists a neighborhood  $V$  of  $\theta_o$  over which  $U_n$  is twice continuously differentiable and  $P_o$ -integrable r.v.  $h$  such that for all  $\alpha \in V$ ,  $U_n^{(2)}(\alpha, x) \leq h(x)$ .

(H2) There exists a sequence  $(a_n) \rightarrow \infty$  such that  $J_n = \text{Var}(\sqrt{a_n} U_n^{(1)}(\theta_o))$  exists and satisfies:

(1) There exists  $J > 0$  such that  $J_n \geq J$  from a certain point on.

(2)  $\sqrt{a_n} J_n^{-1/2} U_n^{(1)}(\theta_o) \longrightarrow N(0, I_p)$ .

(H3) There exists a sequence of nonstochastic matrices  $I_n$  such that  $I_n \geq I > 0$

from a certain point and  $\lim_{n \rightarrow \infty} (U_n^{(2)}(\theta_o) - I_n) = 0$  in  $P_o$  probability.

It is easy to check that the condition (H2) is equivalent to:

(H2') There exists  $(a_n) \rightarrow \infty$  and matrices  $J_n$  satisfying (1) and (2) of the condition (H2)

To prove asymptotic normality of  $\hat{\theta}_T$  we need to verify conditions (H1), (H2') and (H3). Let us denote  $p^{(1)}(x_t, s; \theta)$  the vector of first derivatives of  $p(x_t, s; \theta)$  with respect to  $\theta$ ,

$$u_{t+1}(x_t, s; \theta) = \left[ \frac{X_{t+1}(s)}{1-p} - 1 \right] \times \frac{p^{(1)}}{p}(x_t, s; \theta), \quad \xi_{t+1}(\theta) = \sum_{s \in I(x_t)} u_{t+1}(x_t, s; \theta),$$

$$J_T(\theta_o) = A_T(\theta_o) + B_T(\theta_o) := \frac{1}{n(T)} \sum_{t=0}^{T-1} \text{Var}[\xi_{t+1}(\theta_o) \mid x_t]$$

where  $A_T(\theta_o)$  and  $B_T(\theta_o)$  are given respectively by (6) and (7).

We proceed in two steps: firstly we calculate  $U_T^{(1)}(\theta_o)$ , then  $U_T^{(2)}(\theta_o)$ , the matrix of second derivatives of  $U_T(\theta_o)$ ,  $I_T = A_T(\theta_o)$ ,  $J_T = J_T(\theta_o)$ , the pseudo-information matrices, and show in Lemma 5 the existence of  $I = A(\theta_0) > 0$ ,  $J = J(\theta_0) > 0$  such that  $A_T(\theta_o) \geq A(\theta_0)$ ,  $J_T(\theta_o) \geq \epsilon_o \cdot J(\theta_0)$  for some  $\epsilon_o > 0$ . Secondly, we prove in Lemma 6 a Central Limit



Theorem for  $\sqrt{n(T)}J_T^{-1/2}(\theta_o)U_T^{(1)}(\theta_o)$ . We have :

$$\begin{aligned} U_T^{(1)}(\theta_o) &= \frac{1}{n(T)} \sum_{t=0}^{T-1} \xi_{t+1}(\theta_o), \quad U_T^{(2)}(\theta_o) = A'_T(\theta_o) + C_T(\theta_o), \text{ where :} \\ A'_T(\theta_o) &= \sum_{c \neq (0,0)} \frac{n(T, c)}{n(T)} \times \frac{n_1(T, c)}{n(T, c)} \times \frac{q^{(1)}(c; \theta_o)^t q^{(1)}(c; \theta_o)}{q(c; \theta_o)[1 - q(c; \theta_o)]^2}, \\ C_T(\theta_o) &= \sum_{c \neq (0,0)} \frac{n(T, c)}{n(T)} \left\{ \frac{n_1(T, c)}{n(T, c)} \frac{1}{1 - q(c; \theta_o)} - 1 \right\} (\log q)^{(2)}(c; \theta_o). \end{aligned}$$

Let  $\mathbf{L} = \{L_i, i \in \mathcal{I}\}$  be a partition of  $S$ ,  $\mathcal{L} = \{L_{i(t)}, t \geq 0\}$ , an infinite sequence of elements of  $\mathbf{L}$  and  $(c, 0)$  the configuration of  $\mathcal{N}_2(0)$  such that  $x(\mathcal{N}_1(0)) = c$  and  $x(\mathcal{N}_2(0) \setminus \mathcal{N}_1(0)) \equiv 0$ .

Define

$$\pi_{c,0}^{\mathcal{L}} = \lim_{T \rightarrow \infty} \frac{\sum_{t=0}^{T-1} n^{L_{i(t)}}(x_t, c, 0)}{n(T)}$$

where  $n^L(x_t, c, 0)$  is the number of sites of  $S \cap L$  with configuration  $(c, 0)$ . Let now state the *sub-ergodicity result*:

**Lemma 4** *Let  $\mathcal{C}_1^*$  be the set of configurations on  $\mathcal{N}_1(0)$  such that  $x(0) = 1$ . Then there exists  $\alpha > 0$  and  $\mathcal{L}$ , a sequence of elements of  $\mathbf{L}$ , such that,  $\forall c \in \mathcal{C}_1^*$ , and  $\forall \mathbf{x} \in I_\infty$ , we have  $\pi_{c,0}^{\mathcal{L}} \geq \alpha$ .*

*Proof.* We make use of the the following result : if  $(Y_n, n \geq 1)$  is a sequence of independent Bernoulli variables with parameter  $(q_n, n \geq 1)$  such that  $\beta = \min\{q_n, n \geq 1\} > 0$ , then, a.s.,  $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1, N} Y_n \geq \beta/2$ . The condition  $\beta > 0$  is satisfied if  $\beta_n = \beta(k_n) > 0, n \geq 1$ , where  $k_n \in \mathcal{K} = \{1, \dots, K\}$  is a finite set.

Let  $\mathcal{C}_3^*$  be the set of configurations on  $\mathcal{N}_3(0)$  such that  $x(0) = 1$ . We will define two sequence of instants. The first one,  $0 \leq T_1 < T_2 < \dots < T_{m_T} \leq T - 2$  is the sequence of

informative instants between 0 and  $T - 2$  such that  $x_t(\mathcal{N}_1(s)) \neq (0, 0)$  for at least one site  $s \in S$ . Let us denote  $s_{T_k}$  a site at time  $T_k$  such that  $x_{T_k}(\mathcal{N}_1(s_{T_k})) \neq (0, 0)$ . The second one,  $0 \leq T_1^* < T_2^* < \dots < T_{m_T^*}^* \leq T - 2$ , is the subsequence of the precedent such that  $x_t(s) = 1$  for at least one site of  $S$ ; let us denote  $s_{T_k^*}^*$  an informative site at time  $T_k^*$ . Furthermore, we will choose sites  $s$  and  $s^*$  such that  $s_{T_k} = s_{T_k^*}^*$  when  $T_k = T_k^*$ .

Define  $\mathcal{L}$  as follows: (i) for each informative instant  $t = T_k$ , let  $L_{i(T_k+1)}$  be the subset of  $\mathbf{L}$  containing  $s_{T_k}$ ; (ii) at the others instants,  $L_{i(t)}$  may be chosen arbitrarily in  $\mathbf{L}$ .

For each configuration  $c$  of  $\mathcal{C}_1^*$  and  $k^* = 1, \dots, m_T^*$  let  $Z_{k^*} = \mathbf{1}\{X_{T_{k^*}^*+1}(\mathcal{N}_2(s_{T_{k^*}^*}^*)) = (c, 0)\}$ . As  $P(X_{t+1}(\mathcal{N}_2(s)) \mid x_t)$  depends only on  $x_t(\mathcal{N}_3(s))$ , the variables  $(Z_{k^*}, k^* = 1, \dots, m_T^*)$  are independent Bernoulli variables with parameter

$$q(x_{T_{k^*}^*}(\mathcal{N}_3(s_{T_{k^*}^*}^*)), c) = P(x_{T_{k^*}^*+1}(\mathcal{N}_2(s_{T_{k^*}^*}^*)) = (c, 0) \mid x_{T_{k^*}^*}(\mathcal{N}_3(s_{T_{k^*}^*}^*))).$$

As  $x_{T_{k^*}^*}(s_{T_{k^*}^*}^*) = 1$ , every configuration  $(c, 0)$ ,  $c \in \mathcal{C}_1^*$ , of  $x_{T_{k^*}^*+1}(\mathcal{N}_2(s_{T_{k^*}^*}^*))$  is obtained with positive probability, it follows that:

$$\beta^* = \inf\{q(x_{T_{k^*}^*}(\mathcal{N}_3(s_{T_{k^*}^*}^*)), c) : x_{T_{k^*}^*}(\mathcal{N}_3(s_{T_{k^*}^*}^*)) \in \mathcal{C}_3^*, c \in \mathcal{C}_1^*, T_k^*, \text{ for } k^* \geq 1\} > 0.$$

Let,  $W_k = X_{T_k+1}(s_{T_k})$  for  $k = 1, \dots, m_T$ .  $P(X_{t+1}(s) \mid x_t)$  depending only on  $x_t(\mathcal{N}_1(s))$ , the variables  $(W_k, k = 1, m_T)$  are independent Bernoulli variables with parameter

$$q(x_{T_k}(\mathcal{N}_1(s_{T_k}))) = P(X_{T_k+1}(s_{T_k}) = 1 \mid x_{T_k}(\mathcal{N}_1(s_{T_k}))).$$

Since  $x_t(\mathcal{N}_1(s_{T_k})) \neq (0, 0)$ , the configuration  $X_{T_k+1}(s_{T_k}) = 1$  is obtained with positive probability, that is:

$$\beta = \inf\{q(x_{T_k}(\mathcal{N}_1(s_{T_k}))), x_{T_k}(\mathcal{N}_1(s_{T_k})) \neq (0, 0), T_k, \text{ for } k \geq 1\} > 0.$$

On the one hand, as  $Z_{k^*} = 1$  contributes by 1 in  $n^{L_{i(T_{k^*}^*+1)}}(x_{T_{k^*}^*+1}, c, 0)$ ,

$$\sum_{t=0}^{T-2} n^{L_{i(t+1)}}(x_{t+1}, c, 0) \geq \sum_{k^*=1}^{m_T^*} n^{L_{i(T_{k^*}^*+1)}}(x_{T_{k^*}^*+1}, c, 0) \geq \sum_{k^*=1}^{m_T^*} Z_{k^*}. \quad (8)$$

On the other hand, as there exists at most  $m_T$  informative instants, each with at most  $\sharp(S)$  informative sites,

$$\sum_{t=0}^{T-2} n^{L_{i(t+1)}}(x_{t+1}) \leq \sum_{k=1}^{m_T} n(x_{T_k+1}) \leq \sharp(S) \cdot \frac{m_T}{\sum_{k=1, m_T} W_k} \cdot m_T^*. \quad (9)$$

Furthermore  $m_T^* \geq \sum_{k=1, m_T} W_k$ . It follows from (8) and (9) that:

$$\frac{\sum_{t=0}^{T-2} n^{L_{i(t+1)}}(x_{t+1}, c, 0)}{\sum_{t=0}^{T-2} n^{L_{i(t+1)}}(x_{t+1})} \geq \frac{1}{\sharp(S)} \cdot \frac{1}{m_T^*} \sum_{k=1}^{m_T^*} Z_{k^*} \cdot \frac{1}{m_T} \sum_{k=1}^{m_T} W_k. \quad (10)$$

The choice of  $\mathcal{L}$  entails:

$$\frac{\sum_{t=0}^{T-2} n^{L_{i(t+1)}}(x_{t+1})}{n(T) - n(1)} = \frac{\sum_{t=0}^{T-2} n^{L_{i(t+1)}}(x_{t+1})}{\sum_{t=0}^{T-2} n(x_{t+1})} \geq \frac{1}{\sharp(S)} \quad (11)$$

since  $n^{L_{i(t+1)}}(x_{t+1}) \geq \frac{1}{\sharp(S)} n(x_{t+1})$ , at each informative instant  $t$ , and  $n^{L_{i(t+1)}}(x_{t+1}) = n(x_{t+1}) = 0$ , for the other instants. The result of the Theorem 1 with  $\alpha = \beta \cdot \beta^* / (4 \cdot \{\sharp(S)\}^2)$

is a consequence of (10) and (11) as  $\frac{n(T)-n(1)}{n(T)} \rightarrow 1$  when  $T \rightarrow \infty$ . ■

In the case  $\mathbf{L} = \{S\}$  we obtain the result of Lemma 1 as  $\pi_c \geq \pi_{c,0}^S \geq \alpha$ .

In order to define  $J(\theta_o)$ , let  $L = \{4 \times \mathbb{Z}^2\} \cup \{4 \times \mathbb{Z}^2 + (2, 2)\}$ , and for  $0 \leq k, l \leq 2$ ,  $k + l < 4$  and  $i = 3k + l + 1$ ,  $L_i = \{L + (k, l)\} \cap S$ . Let  $\mathbf{L} = \{L_1, \dots, L_8\}$  the associated partition of  $S$  and  $\mathcal{L} = \{L_{i(t)}, t \geq 0\}$  the sequence defined by Lemma 4. Let  $A(\theta_o)$ ,  $J(\theta_o)$  be the matrices defined by:

$$A(\theta_o) = \sum_{c \neq (0,0)} \pi_c \times \frac{q^{(1) \cdot t} [q^{(1)}]}{q[1-q]}(c; \theta_o), \quad J(\theta_o) = \sum_{c \neq (0,0)} \pi_{c,0}^{\mathcal{L}} \times \frac{q^{(1) \cdot t} [q^{(1)}]}{[1-q]^2 [q]^2}(c; \theta_o).$$

Note that variables  $\{(X_{t+1}(s) \mid x_{t+1}(S_1 \setminus I_0^L(x_t)), x_t), s \in I_0^L(x_t)\}$  are independent where  $I_0^L(x_t) = \{s \in L : x_t(\mathcal{N}_2(s) \setminus \mathcal{N}_1(s)) \equiv 0\}$ . Indeed, for  $A \supseteq S_1 \setminus I_0^L(x_t)$  and  $s \in I_0^L(x_t) \setminus A$ :

$$P[X_{t+1}(s) = 0 \mid x_{t+1}(A), x_t] = \gamma_o^{x_t(s)} \cdot (1 - \lambda_o)^{m(x_{t+1}, \partial s)} \quad (12)$$

$$= P[X_{t+1}(s) = 0 \mid x_{t+1}(S_1 \setminus I_0^L(x_t)), x_t] \quad (13)$$

as  $\partial s \cap I_0^L(x_t) = \emptyset$ .

We have the following result:

**Lemma 5** *For  $\theta_o \in \Theta$  and under the condition **(I)** :*

a)  $A(\theta_o)$  is positive definite and  $\underline{\lim}_{T \rightarrow \infty} A_T(\theta_o) = A(\theta_o)$ ,

b)  $J(\theta_o)$  is positive definite and  $\underline{\lim}_{T \rightarrow \infty} J_T(\theta_o) \geq \epsilon_o \times J(\theta_o)$  for some  $\epsilon_o > 0$ .

*Proof :*

a)  $A(\theta_o)$  is a positive semidefinite matrix. Indeed for each  $\alpha = (\alpha_1, \alpha_2) \in \mathbb{R}^2$ :

$${}^t\alpha A(\theta_o) \alpha = \sum_{c \neq (0,0)} \pi_c q(c; \theta_o) [(1 - q(c; \theta_o))]^{-1} \left( {}^t\alpha (\log q)^{(1)}(c; \theta_o) \right)^2 \geq 0$$

as  $\pi_c > 0$ ,  $0 < q(c; \theta_o) < 1$  for  $c \in \mathcal{C}_1^*$ . It remains to prove that  $\alpha = \mathbf{0}$  when  ${}^t\alpha A(\theta_o) \alpha = 0$ .

Let  $c_a = (u_a, v_a)$  and  $c_b = (u_b, v_b)$  be two linearly independent configurations of  $\mathcal{C}_1^*$  and

suppose that  ${}^t\alpha A(\theta_o) \alpha = 0$ . It follows that  ${}^t\alpha (\log q)^{(1)}(c_a; \theta_o) = {}^t\alpha (\log q)^{(1)}(c_b; \theta_o) = 0$

and  $\alpha_1 = \alpha_2 = 0$ . Finally, as  $\pi_c = \underline{\lim}_{T \rightarrow \infty} \frac{n(T, c)}{n(T)}$ ,  $\underline{\lim}_{T \rightarrow \infty} A_T(\theta_o) = A(\theta_o)$ .

b) We show in a similar way that  $J(\theta_o) > 0$ , since  $\pi_{c,0}^{\mathcal{L}} > 0$  for each  $c \in \mathcal{C}_1^*$ . The proof of  $\underline{\lim}_{T \rightarrow \infty} J_T(\theta_o) \geq \epsilon_o \times J(\theta_o)$  is based upon inequality:

$$Var(X) = E_{\mathcal{G}}(Var(X | \mathcal{G})) + Var_{\mathcal{G}}(E(X | \mathcal{G})) \geq E_{\mathcal{G}}(Var(X | \mathcal{G})) \quad (14)$$

where  $\mathcal{G}$  is some  $\sigma$ -field that we have to choose. This idea was initially proposed by J. Jensen and is used in other estimation contexts: Ising model (Guyon & Künsch, 1992), point processes (Jensen & Künsch, 1994), dynamic of lattice data process (Guyon & Hardouin, 2001).

Applying (14) for  $t = 0, \dots, T-1$  with  $X = (\xi_{t+1}(\theta_o) | x_t)$ ,  $\mathcal{G} = \mathcal{G}_{t+1} = \sigma\{X_{t+1}(s), s \in S_1 \setminus I_0^L(x_t)\}$  and  $L = L_{i(t)}$  and using (12), for  $s \in I_0^L(x_t)$  and  $m(X_{t+1}, \partial s) \geq 1$  it follows that:

$$E_{\mathcal{G}_{t+1}}[Var(X_{t+1}(s) | \mathcal{G}_{t+1}, x_t)] \geq \epsilon_o := \gamma_o(1 - \alpha_o)(1 - \lambda_o)^5 \quad (15)$$

where  $\alpha_o = \gamma_o + (1 - \gamma_o)(1 - \lambda_o)^4$ . Indeed  $Var(X_{t+1}(s) | \mathcal{G}_{t+1}, x_t) \geq \gamma_o(1 - \lambda_o)^5$  as  $\gamma_o \cdot (1 - \lambda_o)^4 \leq P(X_{t+1}(s) = 0 | \mathcal{G}_{t+1}, x_t) \leq 1 - \lambda_o$ . Furthermore,  $P[m(X_{t+1}, \partial s) = 0 |$

$x_t)] \leq \alpha_0$  as  $s \in I(x_t)$  and (15) is a consequence of

$$E_{\mathcal{G}_{t+1}}[Var(X_{t+1}(s) \mid \mathcal{G}_{t+1}, x_t)] \geq \gamma_o(1 - \lambda_o)^5 \cdot \sum_{m \geq 1} P[m(X_{t+1}, \partial s) = m \mid x_t].$$

Then, for  $s \in I_0^L(x_t)$ , (15) implies:

$$E_{\mathcal{G}_{t+1}}[Var(u_{t+1}(x_t, s; \theta_o) \mid \mathcal{G}_{t+1}, x_t)] \geq \epsilon_o \frac{p^{(1)} - {}^t p^{(1)}}{[1 - p]^2 [p]^2}(x_t, s; \theta_o).$$

Finally, since conditional random variables  $\{(X_{t+1}(s) \mid \mathcal{G}_{t+1}, x_t), s \in I_0^L(x_t)\}$  are independent, it follows that:

$$\begin{aligned} E_{\mathcal{G}_{t+1}}[Var(\xi_{t+1}(\theta_o) \mid \mathcal{G}_{t+1}, x_t)] &= E_{\mathcal{G}_{t+1}}\left[\sum_{s \in I_0^L(x_t)} Var(u_{t+1}(x_t, s; \theta_o) \mid \mathcal{G}_{t+1}, x_t)\right] \\ &\geq \epsilon_o \sum_{s \in I_0^L(x_t)} \frac{p^{(1)} - {}^t p^{(1)}}{[1 - p]^2 [p]^2}(x_t, s; \theta_o). \end{aligned}$$

The result follows from the last inequality, and

$$Var[\xi_{t+1}(\theta_o) \mid x_t] \geq \frac{1}{n(T)} \sum_{t=0}^{T-1} E_{\mathcal{G}_{t+1}}[Var(\xi_{t+1}(\theta_o) \mid \mathcal{G}_{t+1}, x_t)] \cdot \blacksquare$$

**Lemma 6** As  $T$  tends to  $\infty$ :  $\sqrt{n(T)} J_T^{-1/2}(\theta_o) U_T^{(1)}(\theta_o) \xrightarrow{d} \mathbf{G}_2(0, I_2)$

*Proof* : We will show that for each  $\alpha \in \mathbb{R}^2$ ,  $\alpha \neq \mathbf{0}$ ,  $\frac{1}{\sqrt{n(T)}} {}^t \alpha J_T^{-1/2}(\theta_o) \sum_{t=0}^{T-1} \xi_{t+1}(\theta_o) \xrightarrow{d} \mathbf{G}_1(0, {}^t \alpha \alpha)$ . We apply a result for triangular martingale arrays (see Hall & Heyde, Theorem 3.2, Dacunha-Castelle & Duflo, Theorem 2.8.43) to  $(S_{T,t}, 1 \leq t \leq T)$ ,  $(\mathcal{F}_{T,t}, 1 \leq t \leq T)$

where:

$$S_{T,t} = \frac{1}{\sqrt{n(T)}} {}^t\alpha J_T^{-1/2}(\theta_o) \sum_{t'=0, t-1} \xi_{t'+1}(\theta_o), \text{ and } \mathcal{F}_{T,t} = \sigma(X_{t'}, 1 \leq t' \leq t) \equiv \mathcal{F}_t.$$

Since for each  $T \geq 0$  and  $1 \leq t \leq T$ ,  $E[\xi_{t+1}(\theta_o) \mid \mathcal{F}_{T,t}] = E[\xi_{t+1}(\theta_o) \mid x_t] = 0$ ,

$\{S_{T,t}, \mathcal{F}_{T,t}, 1 \leq t \leq T\}$  is a zero-mean, square-integrable martingale for each  $T \geq 0$ .

From (??) we get:

$$V_{T,T} = \frac{1}{n(T)} {}^t\alpha J_T^{-1/2}(\theta_o) \sum_{t=0}^{T-1} \text{Var}[\xi_{t+1}(\theta_o) \mid x_t] J_T^{-1/2}(\theta_o) \alpha = {}^t\alpha \alpha > 0.$$

It remains to verify the Lindeberg condition. Let  $B^2 = \max_{i,j=1,2} |b_i \cdot b_j|$  where for  $i = 1, 2$ ,  $b_i = \sup_{\theta_o \in \Theta, c \in \mathcal{C}_1} |\log q_i^{(1)}(c; \theta_o)| \cdot \frac{2-q(c; \theta_o)}{1-q(c; \theta_o)}$ . It follows that:

$$\begin{aligned} & \sum_{t=0}^{T-1} E \left( \frac{1}{n(T)} [{}^t\alpha J_T^{-1/2}(\theta_o) \xi_{t+1}(\theta_o)]^2 \cdot \mathbf{1}_{\left| \frac{1}{\sqrt{n(T)}} {}^t\alpha J_T^{-1/2}(\theta_o) \xi_{t+1}(\theta_o) \right| > \epsilon} \mid x_t \right) \\ & \leq \frac{B^2}{n(T)} {}^t\alpha J_T^{-1}(\theta_o) \alpha \cdot \sum_{t=0}^{T-1} P \left( \left| {}^t\alpha J_T^{-1/2}(\theta_o) \xi_{t+1}(\theta_o) \right| > \epsilon \sqrt{n(T)} \mid x_t \right) \\ & \leq \frac{B^4}{n(T)} [{}^t\alpha J_T^{-1}(\theta_o) \alpha]^2 \cdot \frac{13}{4\epsilon^2}. \end{aligned}$$

Clearly, for a fixed  $\alpha$ ,  ${}^t\alpha J_T^{-1}(\theta_o) \alpha$  is bounded in  $T$ . Indeed, let  $\mu_{J_T}(\theta_o), \mu_J(\theta_o)$  be respectively the smallest eigenvalue of  $J_T(\theta_o)$  and  $J(\theta_o)$ . From Lemma 5  $\underline{\lim}_T J_T(\theta_o) \geq \epsilon_o J(\theta_o) > 0$ . It follows that  $\mu_{J_T}(\theta_o) \geq \epsilon_o \mu_J(\theta_o) > 0$  and:

$$|{}^t\alpha J_T^{-1}(\theta_o) \alpha| \leq \sup_{\alpha} |{}^t\alpha J_T^{-1}(\theta_o) \alpha| = [\mu_{J_T}(\theta_o)]^{-1} \leq [\epsilon_o \mu_J(\theta_o)]^{-1}.$$

Since for each fixed  $\epsilon > 0$ ,  $\frac{B^4}{n(T)} [{}^t\boldsymbol{\alpha} J_T^{-1}(\theta_o) \boldsymbol{\alpha}]^2 \cdot \frac{13}{4\epsilon^2} \rightarrow 0$  as  $T \rightarrow \infty$ , the Lindeberg condition is fulfilled and the proof of Lemma completed. .... ■

We are now able to establish the asymptotic normality of  $\hat{\theta}_T$ . As  $\hat{\theta}_T \xrightarrow{a.s.} \theta_o$ , it is enough to verify conditions (H1), (H2') and (H3) of Theorem 3.4.5 (Guyon, 1995). Let  $a_T = \sqrt{n(T)}$ .

(H1) is satisfied as  $p(x_t, s; \theta)$  is two times continuously differentiable on  $\theta$  and  $\theta_0$  is an interior point of  $\Theta$ , a compact of  $]0, 1[{}^2$ . Furthermore, denoting  $K^2 = \max_{i,j=1,2}(k_{ij})$  where  $k_{i,j} = \max\{\sup_{c \in \mathcal{C}_1, \theta_o \in \Theta} |\frac{\{q^{(1)} \text{ } ^t q^{(1)}(c; \theta_o)\}_{i,j}}{q(c; \theta_o)[1-q(c; \theta_o)]^2}|, \sup_{c \in \mathcal{C}_1, \theta_o \in \Theta} |\frac{(\log q)_{i,j}^{(2)}(c; \theta_o)}{[1-q(c; \theta_o)]}|\}$  for  $i, j = 1, 2$ , it follows that  $|U_T^{(2)}(\theta_o)| = |A'_T(\theta_o) + C_T(\theta_o)| \leq [\sum_{c \neq (0,0)} n_1(T, c)/n(T) \cdot 2 + 1] \cdot K^2$ .

(H2')  $a_T = \sqrt{n(T)} \rightarrow \infty$  since the process survives forever on  $S$ . Lemma 5 establishes the existence of  $J(\theta_o) > 0$  and Lemma 6 the asymptotic normality of  $\sqrt{a_T} J_T^{-1/2}(\theta_o) U_T^{(1)}(\theta_o)$ .

(H3)  $\lim_{T \rightarrow \infty} A_T(\theta_o) = A(\theta_o)$  from Lemma 5. Finally, as  $n_1(T, c)/n(T, c) \xrightarrow{a.s.} 1 - q(c; \theta_o)$  when  $T \rightarrow \infty$ , it follows that  $C_T(\theta_o) \xrightarrow{a.s.} 0$  and  $A'_T(\theta_o) - A_T(\theta_o) \xrightarrow{a.s.} 0$  and this completes the proof. .... ■

## Acknowledgments

The authors would like to thank the editor and the referees for their helpful comments.

## References

- Besag, J., 1974, Spatial Interaction and the Statistical Analysis of Lattice Systems, *JRSS B*, 36, 192-225.



- Dacunha-Castelle, D., Duflo, M., 1986, *Probability & Statistics* - 2nd vol, Springer.
- Durrett, R., Levin, S.A., 1994, Stochastic spatial models: a user's guide to ecological applications, *Phil. Trans. R. Soc. Lond.*, B 343, 329-350.
- Durrett, R., 1995, *Ten lectures on particle systems, Cours de Saint Flour (1993)*, L.N.M. 1608, Springer, 97-201.
- Fiocco, M., 1997, *Statistical estimation for the supercritical contact process*, PHD Thesis, Leiden University, Mathematical Institute, The Netherlands.
- Fiocco, M., Zwet, W.R., 2003, Parameter estimation for the supercritical contact process, *Bernoulli*, 9, 1071–1092.
- Griffeath, D., 1981, The basic contact processes, *Stoch. Proc. Appl.*, 11, 151-185.
- Guyon, X., 1995, *Random fields on a network: modelling, statistics and applications*, Springer, Berlin.
- Guyon, X., Künsch, H.R., 1992, Asymptotic comparison of estimators in the Ising model, L.N.S. 74, 177-198.
- Guyon, X., Hardouin, C., 2001, Markov chain Markov field dynamics: Models and Statistics, *Statistics*, 35, 593-627.
- Hall, P., Heyde, C.C., 1980, *Martingale limit theory and its application*, Academic Press, New York.
- Jensen, J.L., Künsch, H.R., 1994, On asymptotic normality of pseudo-likelihood estimate for pairwise interaction processes, *Ann. Inst. Statist. Math.*, 46, 475-486.

Mollison, D., (1977) Spatial contact models for ecological and epidemic spread, with discussion, *JRSS B*, 39, 283-326.

Whittle, P., 1954, On stationary process in the plane, *Biometrika*, 41, 434-449.